

09/849,503

Express Mail No. EL897870358US



Europäisches  
Patentamt

European  
Patent Office

Office européen  
des brevets

#15



Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-  
gen stimmen mit der  
ursprünglich eingereichten  
Fassung der auf dem näch-  
sten Blatt bezeichneten  
europäischen Patentanmel-  
dung überein.

The attached documents  
are exact copies of the  
European patent application  
described on the following  
page, as originally filed.

Les documents fixés à  
cette attestation sont  
conformes à la version  
initialement déposée de  
la demande de brevet  
européen spécifiée à la  
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

00830332.3

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN  
THE HAGUE, 25/04/01  
LA HAYE, LE



**THIS PAGE BLANK (USPTO)**



Europäisches  
Patentamt

European  
Patent Office

Office européen  
des brevets

**Blatt 2 der Bescheinigung**  
**Sheet 2 of the certificate**  
**Page 2 de l'attestation**

Anmeldung Nr.:  
Application no.:  
Demande n°: 00830332.3

Anmeldetag:  
Date of filing: 05/05/00  
Date de dépôt:

Anmelder:  
Applicant(s):  
Demandeur(s):  
STMicroelectronics S.r.l.  
20041 Agrate Brianza (Milano)  
ITALY

Bezeichnung der Erfindung:  
Title of the invention:  
Titre de l'invention:  
Motion estimation process and system.

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:  
State:  
Pays:

Tag:  
Date:  
Date:

Aktenzeichen:  
File no.  
Numéro de dépôt:

Internationale Patentklassifikation:  
International Patent classification:  
Classification internationale des brevets:  
H04N7/26, H04N7/36, H04N5/14

Am Anmeldetag benannte Vertragsstaaten:  
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE/UK  
Etats contractants désignés lors du dépôt:

Bemerkungen:  
Remarks:  
Remarques:

**THIS PAGE BLANK (USPTO)**

**"Procedimento e sistema per la stima del movimento"**

Campo dell'invenzione

La presente invenzione si riferisce alle tecniche  
5 di stima del movimento utilizzabili, ad esempio, nei sistemi per il trattamento e la trasmissione di segnali video.

Descrizione della tecnica nota

La maggior parte delle tecniche per la stima del  
10 movimento sviluppate nel corso degli ultimi anni possono essere sostanzialmente ricondotte ad alcuni filoni fondamentali, ossia:

- i) le tecniche di ricerca rapida, che cercano di alleggerire la complessità riducendo il numero dei  
15 vettori di moto candidati che vengono testati nell'area di ricerca. La scelta dei vettori viene pilotata in modo euristico in maniera da cercare di localizzare il minimo assoluto della funzione di costo: esemplificative di tale filone sono le soluzioni note  
20 come Three Step Search, Fast Search, New Three Step Search e Four Step Search documentate, ad esempio, dai lavori B. Furth, J. Greenberg, R. Westwater, "Motion Estimations Algorithms for Video Compression", Kluwer Academic Publishers, 1997; R. Li, B. Zeng, M. L. Liou,  
25 "A new Three-Step Search Algorithm for Block Motion Estimation", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 4, No. 4, Aug. 1994, pp. 438-441; L. -M. Po, W. -C. Ma, "A novel Four-Step Search Algorithm for Fast Block Motion Estimation",  
30 IEEE Transactions on Circuits and Systems for Video Technology, Vol. 6, No. 3, Jun. 1996, pp. 313-317;

- ii) le tecniche di adattamento (matching) veloce, in cui la complessità viene ridotta diminuendo il numero di punti su cui viene calcolata la funzione  
35 di costo; esemplificative di tale filone sono le

tecniche di sottocampionamento a quinconce e di trasformata su un bit descritte nei lavori K. Lengwehasatit, A. Ortega, "A Novel Computationally Scalable Algorithm for Motion Estimation", VCIP '98, Jan. 1998; B. Natarajan, V. Bhaskaran, K. Konstantinides, "Low-Complexity Block-Based Motion Estimation via One-Bit Transforms", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 7, Aug. 1997, pp. 702-706;

10       iii) le tecniche che utilizzano una correlazione spazio-temporale: qui il vettore di moto viene scelto fra quelli che sono già stati calcolati nell'estremità corrente per le trame o frame precedenti. Spesso questo vettore è il punto di partenza di un processo di  
15       affinamento, così come documentato dal lavoro di J. Chalidabhongse, C. -C. Kuo, "Fast Motion Vector Estimation Using Multiresolution-Spatio-Temporal Correlations", IEEE Transaction on Circuits and Systems Technology, Vol.7, No. 3, June 1997, pp. 477-488 e dal  
20       lavoro di K. Lengwehasatit e A. Ortega già citato in precedenza;

      iv) le tecniche gerarchiche o a risoluzione multipla, in cui i vettori di moto vengono dapprima ricercati in un'immagine a bassa risoluzione e poi  
25       affinati nell'immagine a risoluzione normale, così come documentato nel lavoro di J. Chalidabhongse e C. -C. Kuo già citato in precedenza; e

      v) le tecniche ottimizzate in funzione della velocità-distorsione (Rate-Distortion optimized  
30       techniques) che non costituiscono veri e propri algoritmi a bassa complessità, dal momento che la funzione di costo è in qualche misura più complicata, prendendo in conto non solo l'energia del segnale differenziale compensato nei confronti del movimento ma  
35       anche quella che deriva dalla scelta del vettore di

moto (il numero di bit necessari per codificarlo, i modi di codifica, ecc.), così come documentato dal lavoro di F. Kossentini, Y. -W. Lee, Mark J.T. Smith, R.K. Ward, "Predictive RD Optimized Motion Estimation  
5 for Very Low Bit-Rate Video Coding", IEEE Journal on Selected Areas in Communications, Vol. 15, No. 9, Dec. 1997, pp. 1752-1763.

Scopi e sintesi della presente invenzione

La presente invenzione si prefigge lo scopo di  
10 fornire una soluzione che, conservando caratteristiche di bassa complessità, consenta di conseguire risultati sostanzialmente confrontabili con i risultati ottenibili con le tecniche basate su una ricerca esaustiva.

15 Secondo la presente invenzione, tale scopo viene raggiunto grazie ad un procedimento avente le caratteristiche richiamate in modo specifico nelle rivendicazioni che seguono. L'invenzione riguarda anche il relativo sistema.

20 La soluzione secondo l'invenzione risulta radicalmente diversa dalle tecniche di ricerca rapida a cui si è fatto cenno in precedenza, in cui la complessità computazionale viene contenuta limitando in modo corrispondente l'area di ricerca ad un ristretto  
25 numero di punti.

La soluzione secondo l'invenzione differisce anche dalle soluzioni denominate Rate-Distorsion Optimized a cui si è fatto riferimento in precedente, in cui, per valutare i vettori di moto, viene utilizzata una  
30 funzione lagrangiana che combina l'energia residua con il numero di bit consumati. Un vettore di moto viene stimato a partire da un'area di ricerca a rombo divisa in livelli sulla base della correlazione esistente fra macroblocchi adiacenti: a partire da tutti i  
35 macroblocchi dello stesso livello vengono considerati i

vettori di moto che vengono quindi combinati in modo opportuno con una funzione di predizione (valore medio, valore mediano, valore medio ponderato, valore medio statistico) per generare un nuovo vettore di moto sul  
5 quale valutare la funzione lagrangiana e decidere se devono essere testati altri livelli.

Rispetto alle soluzioni descritte nei lavori di Chalidabhongse e Kuo e di Lengwehasatit e Ortega citati in precedenza, la soluzione secondo l'invenzione  
10 differisce essenzialmente durante la fase di affinamento dal momento che la scelta del candidato migliore, sulla base della correlazione spazio-temporale fra i vettori rispetto ai macroblocchi adiacenti risulta la stessa per tutti gli algoritmi.  
15 Tuttavia, in questi algoritmi la fase di affinamento è un processo ricorsivo in quanto la griglia di affinamento viene spostata fino a quando non si ottiene una riduzione della funzione di costo.

Al contrario, nella soluzione secondo  
20 l'invenzione, la fase di affinamento viene condotta su una griglia con un numero di punti che è fisso. La griglia stessa risulta amplificabile in modo da poter recuperare gli errori di stima, ma il numero di punti della griglia è sempre lo stesso.

25 In sostanza, la soluzione secondo l'invenzione presenta, nei confronti delle soluzioni precedentemente note, tre vantaggi fondamentali:

- si tratta di una soluzione a complessità costante, in quanto il numero delle operazioni è lo  
30 stesso per tutti i macroblocchi,

- il campo dei vettori di moto è indipendente da qualunque finestra di ricerca, e

- si tratta di una soluzione a bassa complessità in quanto il numero delle operazioni richieste risulta  
35 trascurabile rispetto ad altri algoritmi.



Descrizione particolareggiata dei disegni annessi

L'invenzione verrà ora descritta, a puro titolo di esempio non limitativo, con riferimento ai disegni annessi, nei quali:

- 5           - la figura 1 illustra, sotto forma di uno schema a blocchi, la struttura di uno stadio di un sistema per l'elaborazione di segnale video in cui può trovare applicazione la soluzione secondo l'invenzione,
- la figura 2 illustra i criteri di svolgimento  
10 della funzione di correlazione spaziale e temporale su macroblocchi secondo l'invenzione,
- la figura 3 illustra lo svolgimento della funzione di affinamento nell'ambito della soluzione secondo l'invenzione, e
- 15           - le figure 4, 5 illustrano in modo schematico l'applicazione dell'invenzione allo standard video H.263+.

Descrizione particolareggiata di un modo preferito di attuazione dell'invenzione

- 20           Nello schema della figura 1 il riferimento numerico 1 indica nel complesso uno stadio per l'elaborazione di segnali video digitali destinato ad essere associato, ad esempio, ad un codificatore di segnale video.
- 25           Nello schema della figura 1 (da ritenersi complessivamente noto) il riferimento 10 indica la linea di ingresso su cui viene alimentato il segnale video in forma numerica. Il riferimento 11 indica un blocco di stima del movimento destinato ad operare sul  
30 segnale in arrivo sulla linea 10 in funzione dei dati di predizione, provenienti su una linea 12 a partire da una memoria 13 in cui sono immagazzinate le trame o frame di predizione. La linea di uscita del blocco 11, indicata con 14, è quella su cui sono presenti i  
35 vettori di stima del moto.

Oltre che alla memoria 13, la linea 14 fa anche capo al complesso delle linee di uscita del dispositivo 1, complessivamente indicate con 15.

Il riferimento 16 indica il blocco di controllo  
5 della codifica destinato ad agire, fra l'altro, su due blocchi commutatori 17a, 17b.

Il blocco commutatore 17a è destinato a trasferire verso il blocco 18, che implementa la funzione correntemente denominata DCT (acronimo per Discrete  
10 Cosine Transform ..... ) alternativamente:

- il segnale presente sulla linea di ingresso 10,
- il segnale ottenuto in un blocco di differenza 19 come differenza fra il segnale presente sulla linea di ingresso 10 ed il segnale prelevato dalla linea di  
15 uscita 12 della memoria 13.

L'uscita del blocco 18 viene inviata verso un blocco di quantizzazione 20 che agisce sotto il controllo del blocco di controllo di codifica 16.

L'uscita del blocco 20, indicata con 21,  
20 costituisce un'altra delle uscite 15 del dispositivo 1. Il segnale presente sulla linea 21 viene anche inviato verso due blocchi 22, 23 che svolgono, in cascata, funzioni inverse rispetto alle funzioni svolte dai blocchi 20 e 18 in modo da ricostituire campioni  
25 suscettibili di essere trasmessi verso la memoria 13 attraverso un'ulteriore linea 24. Nella linea 24 è inserito un nodo di somma 25 a cui fa capo la linea di uscita del commutatore 17b. La funzione di quest'ultimo è coordinata con la funzione del commutatore 17a in  
30 modo tale per cui quando il blocco 17a trasferisce verso il blocco 18 il segnale presente sulla linea di ingresso 10, nel blocco 25 al segnale in uscita dal blocco 23 viene sommato il segnale presente sulla linea 12.

In aggiunta alle linee 14 e 21 già descritte in precedenza, il complesso delle linee di uscita 15 del dispositivo 1 comprende ancora ulteriori linee 26, 27 e 28 tutte provenienti dal blocco di controllo della  
5 codifica 16 e destinate a convogliare i rispettivi segnali di flag intra/inter, di flag indicativo dell'avvenuta trasmissione ovvero della mancata trasmissione e di indicazione relativa alla quantizzazione. Si apprezzerà che la linea 28 coincide  
10 di fatto con la linea di pilotaggio del blocco 20 da parte del blocco 16.

Tutto quanto detto in precedenza corrisponde peraltro a criteri di per sé complessivamente noti, qui richiamati più che altro al fine di fornire  
15 un'illustrazione di un possibile contesto di applicazione dell'invenzione.

In generale, un qualsiasi procedimento di stima del moto (così come implementato, ad esempio, dal blocco 11 della figura 1) deve associare a ciascun  
20 macroblocco in cui è stata divisa una trama o frame di una sequenza di immagini video, un vettore di moto che punta al macroblocco di un precedente frame di riferimento quanto più simile possibile al macroblocco corrente. A questo fine esso deve confrontare con il  
25 blocco corrente alcuni o tutti i macroblocchi del frame di riferimento nell'ambito di un'area di ricerca limitata in modo opportuno tramite una funzione di costo adatta che misura la differenza fra i due macroblocchi.

30 Come noto, ci sono diverse funzioni che possono stimare l'adattamento e/o l'errore residuo. Si tratta in generale di funzioni della differenza per punti fra i due macroblocchi.

Nella forma di attuazione dell'invenzione al  
35 momento preferita, viene utilizzata la funzione

denominata SAD (acronimo per Sum of Absolute Differences ovvero Somma delle Differenze Assolute). Si tratta di una funzione tale da offrire un buon compromesso fra la precisione e complessità di calcolo.

- 5 Come possibile alternativa si potrebbe utilizzare, ad esempio, la funzione SSD (acronimo di Sum of Squared Differences ovvero Somma delle Differenze al quadrato). Questa soluzione alternativa presenta qualche vantaggio in termini di affidabilità, dal  
10 momento che calcola l'energia effettiva associata alla differenza. Presenta tuttavia l'inconveniente legato al fatto di chiedere una moltiplicazione.

La funzione SAD è definita nel modo seguente:

$$\text{SAD}(m,n) = \sum |B(i,j) - B'(i+m,j+n)| \quad (I)$$

- 15 dove la sommatoria si estende a tutti i valori  $i,j$  appartenenti al macroblocco e  $B(i,j)$  è l'intensità di pixel nella posizione  $i,j$  del frame corrente e  $B'(i+m,j+n)$  è l'intensità di pixel corrispondente nel frame precedente, spostato di un vettore con componenti  
20  $x,y$  pari rispettivamente a  $m$  e  $n$ .

La funzione SAD è una misura della qualità del vettore di moto, dal momento che un vettore di moto corretto produce una SAD ridotta, mentre un vettore di moto erroneo produce una SAD di valore elevato.

- 25 Pertanto il calcolo della SAD viene eseguito per tutti i candidati suscettibili di essere scelti come vettori di moto per un macroblocco dato. Il vettore di moto tale da produrre la SAD più bassa viene scelto come vettore del moto per il macroblocco in questione e  
30 viene utilizzato per la funzione di predizione relativa.

Il procedimento di stima del moto secondo l'invenzione appartiene alla classe dei procedimenti basati sulla valutazione della correlazione spazio-

temporale esistente fra vettori di moto appartenenti a blocchi adiacenti.

Se è vera l'ipotesi che il campo di moto varia lentamente tanto localmente quanto da frame a frame, allora è possibile sottoporre a controllo soltanto piccoli insiemi di vettori di moto candidati. Il vettore di moto candidato che produce la SAD minima viene scelto come vettore di moto predittore che, dopo una fase di affinamento, dà origine al vettore di moto finale.

Dal momento che prevede l'impiego di un vettore di moto con funzione di predittore, questa soluzione corrisponde pertanto ad un procedimento di predizione.

Come tutti i procedimenti di questo tipo, esso si basa essenzialmente su due fasi, ossia:

- l'identificazione dei predittori candidati, e
- l'affinamento del predittore migliore.

Alla fine dell'intero procedimento, a ciascun macroblocco viene associato il vettore di moto che produce la SAD più bassa.

La prima fase richiamata in precedenza consiste quindi nell'identificare, nell'ambito di un insieme di candidati, il vettore di moto predittore identificato come il migliore e quindi come quello a cui applicare la successiva fase di affinamento.

Nell'intento di perseguire una soluzione a bassa complessità, l'invenzione mira a ridurre quanto più possibile il numero di candidati e, allo stesso modo, a scegliere candidati "buoni", presumibilmente vicini al vettore giusto.

Nella soluzione secondo l'invenzione, il criterio per costituire l'insieme dei candidati è quello di scegliere i vettori di quella posizione relativa di macroblocco, nel frame corrente e nel frame precedente, prossima al vettore corrente. Naturalmente c'è la

limitazione data dal fatto che un vettore predittore deve essere calcolato in precedenza. Poiché i macroblocchi vengono normalmente sottoposti a scansione in ordine lessicografico, per il frame corrente sono  
5 disponibili come vettori utilizzabili come predittori solo i vettori appartenenti a macroblocchi disposti al di sopra ed a sinistra di quello corrente.

La figura 2 fa vedere una soluzione che consente di ridurre il numero di candidati mantenendo basso  
10 l'onere computazionale. In una forma di attuazione preferita, la soluzione secondo l'invenzione utilizza - per un macroblocco - quattro predittori candidati; si tratta per la precisione di due predittori spaziali (rispettivamente a sinistra ed al di sopra del  
15 macroblocco corrente) presi dallo stesso frame e due predittori temporali (rispettivamente omologo ed a sinistra rispetto al macroblocco corrente) nel frame precedente.

Questa soluzione è schematicamente illustrata  
20 nella figura 2, dove a sinistra è rappresentato il frame precedente (frame  $t-1$ ) ed a destra il frame corrente (frame  $t$ ).

Il macroblocco corrente è quello lasciato in bianco nella parte di sinistra della figura 2 ed il  
25 macroblocco G nella parte di destra della stessa figura 2 rappresenta il macroblocco omologo del frame precedente.

Per chiarezza, i due predittori spaziali presi dal frame  $t$  sono quelli indicati dalle lettere A e B. I due  
30 predittori temporali sono quelli indicati dalle lettere G ed H nel frame  $t-1$ .

Il vettore di moto che, nell'insieme così definito, produce l'errore residuo più basso (ad esempio la funzione SAD di valore più basso) viene  
35 utilizzato come punto di partenza della successiva fase

di affinamento e viene indicato nel seguito quale "migliore predittore del vettore di moto".

Una volta determinato il predittore in questione, il procedimento continua con una fase di affinamento  
5 attuata tramite una griglia di  $n$  punti che viene applicata allo stesso.

Un esempio di tale griglia è rappresentato nella figura 3. La griglia in questione è formata da quattro punti I a IV ai vertici di un quadrato e quattro  
10 ulteriori punti V a VIII situati ai vertici di un quadrato di dimensioni più piccole con i lati inclinati di  $45^\circ$  rispetto ai lati del quadrato ai cui vertici si trovano i punti I a IV.

La quadrettatura di fondo rappresentata nella  
15 figura 3 corrisponde ad una griglia su mezzo pixel (half pixel grid) e si è supposto che i punti I a VIII siano disposti in corrispondenza di punti di incrocio delle linee di tale quadrettatura.

La suddetta griglia è da considerarsi centrata  
20 intorno alla posizione verso cui punta il migliore vettore di moto MV.

I punti I a VIII presentano, rispetto al centro, una distanza definita da una funzione lineare che dipende dall'errore di adattamento riscontrato durante  
25 il passo precedente (ad esempio, una distanza definita come prodotto fra rispettivi coefficienti e le rispettive SAD). Inoltre, dal momento che il procedimento opera con vettori di moto con la precisione di mezzo pixel, in una forma di attuazione  
30 preferita, i punti presentano una distanza di mezzo pixel dal centro.

Dopo aver considerato tutti i punti della griglia il vettore con la SAD più bassa viene scelto come vettore di moto per il macroblocco corrente e viene  
35 utilizzato per la sua compensazione del moto.

In quei casi in cui il miglior predittore può essere sbagliato, ad esempio in presenza di un cambiamento di scena nell'ambito di un movimento a rapida evoluzione, la correzione della griglia può  
5 essere amplificata. La griglia viene amplificata soltanto quando il miglior predittore presenta un valore di SAD elevato, il che significa che probabilmente non è un buon predittore.

A partire dall'analisi delle distribuzioni di SAD  
10 ottenute per alcune sequenze campione, è possibile fissare una funzione lineare della SAD (ad esempio la stessa funzione coefficiente x SAD vista in precedenza) che consente di calcolare l'entità della funzione di amplificazione ed il numero dei punti, discriminando  
15 altresì le condizioni in cui è necessario procedere all'amplificazione rispetto a quelle in cui tale passo non risulta necessario.

Una misura della complessità della soluzione secondo l'invenzione (in particolare per quanto  
20 riguarda il confronto con altre soluzioni già note, quali quelle basate su un algoritmo di ricerca completa) è dato dal numero di volte in cui si procede al calcolo della funzione di costo. Nel caso specifico, la misura può essere identificata nel numero di  
25 operazioni di calcolo della funzione SAD per macroblocco.

Va notato che questa misura di complessità risulta del tutto indipendente dal formato e dalla frequenza di frame della sequenza su cui viene attuato il  
30 procedimento di codifica ed è pertanto valida per tutti gli algoritmi di codifica.

Per un algoritmo di ricerca completo, il numero di calcoli della SAD dipende dalle dimensioni dell'area di ricerca. Ad esempio, nel modo di funzionamento in  
35 predizione di default dallo standard video H.263+ (si



veda al riguardo il documento Image Processing Lab, University of British Columbia, "TMN (H.263+) encoder/decoder, version 3.0", TMN (H.263+) codec, September 1997) l'area di ricerca presenta un'ampiezza di 32x32 pixel. Inoltre, il miglior vettore di moto che viene ritrovato viene affinato con una precisione di mezzo pixel con l'applicazione di una griglia su n punti. Pertanto, la complessità computazionale dell'algoritmo di ricerca completa è

10  $OP_{MB} = \text{Num.SAD}_{\text{integer-pixel}} + \text{Num.SAD}_{\text{half-pixel}} = 32 \times 32 + 8 = 1032$   
(II)

Al contrario, la soluzione secondo l'invenzione richiede, nell'esempio di attuazione qui illustrato, quattro calcoli di SAD per la scelta del predittore migliore (vedere la figura 2) ed otto calcoli per la griglia di affinamento (vedere la figura 3). A questi calcoli bisogna aggiungere il calcolo della funzione SAD per il vettore nullo. In conclusione, il numero complessivo di calcoli della funzione SAD è dato, nella soluzione secondo l'invenzione, da:

20  $OP_{MB} = \text{Num.SAD}_{\text{half-pixel}} + \text{Num.SAD}_{\text{Null-Vector}} = 12 + 1 = 13$   
(III)

La soluzione secondo l'invenzione consente quindi di ridurre di circa il 90% l'onere computazionale rispetto all'algoritmo di ricerca completa. Inoltre, il costo è stabile poiché il numero totale di calcoli della funzione SAD è lo stesso per ciascun macroblocco. Da ciò deriva un notevole vantaggio della soluzione secondo l'invenzione rispetto ad altri algoritmi di stima del movimento con complessità variabile derivante dalla loro forma interattiva.

La soluzione secondo l'invenzione si presta a sviluppi particolarmente vantaggiosi con riferimento allo standard video H.263+ già richiamato in precedenza.

Tale standard prevede alcun modi opzionali che consentono di migliorare la qualità soggettiva delle sequenze ricostruite. In generale, però, l'impiego di questi modi produce un incremento più o meno  
5 significativo dell'onere computazionale.

Nel contesto dell'illustrazione della presente invenzione è vantaggioso considerare, fra tali opzioni comprese nello standard, alcune opzioni che risultano più vicine al problema della stima del moto, ossia il  
10 modo con frame PB migliorato (Improved PB-frames Mode, brevemente indicato come modo IPB) ed il modo di predizione avanzata (correntemente denominato APM, acronimo per Advanced Prediction Mode). L'integrazione di queste opzioni in un algoritmo predittivo le rende  
15 più omogenee con lo standard.

Per quanto riguarda la descrizione del modo IPB, oltre che al documento generale inerente allo standard H.263+ già citato in precedenza, si può far utilmente riferimento ai documenti ITU Telecom Standardization  
20 Sector of ITU, "Video Coding for Low Bitrate Communication", Draft 21 ITU-T, Recommendation H.263+ Version 2, Jan. 1998 e ITU Telecom Standardization Sector of ITU, "Video Codec Test Model, Near Term, Version 10", TMN10 ITU-T, Apr. 1998.

25 Come evoluzione rispetto allo standard, è possibile introdurre un solo cambiamento al fine di consentire la sua integrazione in modo funzionale nell'algoritmo. Questo cambiamento riguarda l'ordine in cui si realizza la stima del movimento. Lo standard  
30 applica infatti la stima del movimento dapprima sul frame P della coppia PB e poi sul frame B della coppia (questo anche se l'ordine temporale è esattamente contrario).

Questa soluzione presenta però l'inconveniente  
35 dato dal fatto che i predittori temporali del frame B

della coppia PB verrebbero scelti a partire dal campo di moto di riferimento relativo al frame successivo dal punto di vista temporale.

Nella soluzione prospettata nella figura 4 si vede  
5 invece che, in una possibile applicazione della soluzione secondo l'invenzione, l'ordine di stima viene invertito così da rispettare la continuità temporale dei frame. In particolare, la figura 4 fa vedere, nell'ambito della sequenza di frame nel modo IPB, la  
10 gestione della prima coppia PB (ossia i frame  $P_2B_2$ ).

Gli stessi documenti ITU citati in precedenza forniscono la descrizione particolareggiata del modo APM.

In questo contesto applicativo, la soluzione  
15 secondo l'invenzione offre, ad esempio, la possibilità di associare al frame di un macroblocco quattro vettori invece di uno solo. Tutto questo in modo da poter realizzare una stima più accurata.

Così come indicato nello standard, l'impiego  
20 dell'APM consiste in un'applicazione dell'algoritmo di stima del movimento al macroblocco corrente (così da ottenere un vettore del moto associato come nel modo base) ed in una successiva applicazione dello stesso algoritmo per ciascun blocco  $8 \times 8$  che costituisce il  
25 macroblocco  $16 \times 16$  puntato dal vettore ritrovato in precedenza (così da ottenere quattro vettori associati per il macroblocco corrente). Tramite particolari regole di preferenza, il codificatore sceglie il modo ottimale per realizzare la predizione del macroblocco  
30 corrente.

La soluzione schematicamente rappresentata nella figura 5, basata sull'applicazione della soluzione secondo l'invenzione, prevede di applicare il  
procedimento predittivo nella sua versione completa  
35 soltanto al vettore di ricerca associato al macroblocco

16x16. Al contrario, soltanto la fase di affinamento dell'algoritmo viene applicata alla ricerca dei quattro vettori per i blocchi 8x8.

Questo schema è rappresentato nella figura 5 dove  
5 CM indica il macroblocco corrente e PM il macroblocco ottenuto tramite predizione.

La freccia BV indica l'ottenimento del vettore di moto migliore per il macroblocco 16x16 corrente attuata tramite il procedimento predittivo, mentre le frecce  
10 MV<sub>1</sub>, MV<sub>2</sub>, MV<sub>3</sub> e MV<sub>4</sub> illustrano l'ottenimento di quattro rispettivi vettori migliori ottenuti, attraverso la fase di affinamento, per i rispettivi quattro blocchi 8x8 del macroblocco PM ottenuto tramite predizione.

Nel caso del modo APM, la complessità del  
15 procedimento secondo l'invenzione risulta maggiore se confrontata la versione di base in considerazione del fatto che il numero dei candidati predittori è più elevato. Ad esempio, se l'algoritmo usa l'equivalente di ventun predittori per un totale di ventuno  
20 operazioni di calcolo della SAD per macroblocco (si sottolinea che questo è il numero massimo delle operazioni), il guadagno in termini di complessità è 1,62 volte rispetto a quello della versione base.

Naturalmente, fermo restando il principio  
25 dell'invenzione, i particolari di realizzazione e le forme di attuazione potranno essere ampiamente variati rispetto a quanto descritto ed illustrato, senza per questo uscire dall'ambito della presente invenzione così come definita dalle rivendicazioni annesse.

### RIVENDICAZIONI

1. Procedimento per la stima del movimento in segnali video (10) organizzati in frame successivi suddivisi in macroblocchi tramite l'identificazione di vettori di moto, il procedimento comprendendo una prima fase di identificazione in cui, a partire da un vettore di moto corrente, un predittore di vettore di moto migliore viene identificato nell'ambito di un insieme di candidati ed una seconda fase di affinamento del predittore migliore così identificato; detto insieme di candidati essendo formato a partire da vettori appartenenti a macroblocchi prossimi a detto vettore corrente nell'ambito del frame corrente e del frame precedente.

2. Procedimento secondo la rivendicazione 1, caratterizzato dal fatto che detto insieme include, nel frame precedente, il vettore omologo a detto vettore di moto corrente.

3. Procedimento secondo la rivendicazione 1 o la rivendicazione 2, caratterizzato dal fatto che comprende l'operazione di sottoporre a scansione detti macroblocchi in ordine lessicografico nell'ambito di ciascun frame, per cui detto insieme include, nell'ambito del frame corrente, vettori appartenenti a macroblocchi disposti al di sopra ed a sinistra del macroblocco corrente.

4. Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che detto insieme include vettori appartenenti ad almeno uno dei macroblocchi scelti nel gruppo costituito da:

- nel frame corrente, macroblocchi situati a sinistra ed al di sopra del macroblocco corrente, e
- nel frame precedente, il macroblocco omologo e i macroblocchi situati a sinistra del macroblocco corrente.

5. Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che detto predittore migliore viene identificato, nell'ambito di detto insieme, quale predittore che  
5 minimizza una funzione di misura di errore residuo.

6. Procedimento secondo la rivendicazione 5, caratterizzato dal fatto che detta funzione è la funzione SAD (somma delle differenze assolute).

7. Procedimento secondo la rivendicazione 5,  
10 caratterizzato dal fatto che detta funzione è la funzione SSD (somma delle differenze al quadrato).

8. Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che detta fase di affinamento comprende l'operazione di  
15 formare una griglia di n punti centrati intorno alla posizione centrale puntata dal vettore di modo migliore (MV), i punti di detta griglia presentando una distanza rispetto a detto centro definita quale funzione, preferibilmente lineare, dell'errore di adattamento  
20 definito in detta prima fase di identificazione.

9. Procedimento secondo la rivendicazione 8, caratterizzato dal fatto che detta distanza è definita come prodotto fra un coefficiente e detta funzione di errore.

25 10. Procedimento secondo la rivendicazione 8 o la rivendicazione 9, caratterizzato dal fatto che detta funzione di errore è la funzione SAD (somma delle differenze assolute).

11. Procedimento secondo una qualsiasi delle  
30 rivendicazioni 8 a 10, caratterizzato dal fatto che comprende l'operazione di scegliere, come vettore di moto per il macroblocco corrente, nell'ambito di detta griglia, il vettore che minimizza un'ulteriore funzione di errore residuo.

12. Procedimento secondo la rivendicazione 11, caratterizzato dal fatto che detta ulteriore funzione di errore residuo è la funzione SAD (somma delle differenze assolute).

5        13. Procedimento secondo una qualsiasi delle rivendicazioni 8 a 12, caratterizzato dal fatto che comprende l'operazione di amplificare selettivamente detta griglia per un fattore di amplificazione dato.

10       14. Procedimento secondo la rivendicazione 13, caratterizzato dal fatto che comprende l'operazione di monitorare la qualità di detto predittore in funzione di una funzione di errore data e dal fatto che detta griglia viene amplificata quando detto predittore presenta una funzione di errore superiore a detto  
15       valore dato.

15       15. Procedimento secondo la rivendicazione 13 o la rivendicazione 14, caratterizzato dal fatto che detto fattore di amplificazione della griglia viene calcolato in funzione dell'analisi della distribuzione di detta  
20       funzione di errore per alcune sequenze campione.

25       16. Procedimento secondo una qualsiasi delle rivendicazioni 1 a 15, applicato al modo di funzionamento IPB dello standard video H.236+ per il trattamento di coppie di frame PB, comprendente  
25       l'operazione di scegliere i predittori temporali del frame B di una coppia BP prima dei predittori temporali del frame P della coppia BP stessa.

30       17. Procedimento secondo una qualsiasi delle rivendicazioni 1 a 15, applicato al modo di funzionamento APM dello standard video H.263+, in cui  
30       al frame di un macroblocco vengono associati una pluralità di vettori, preferibilmente in numero di quattro, associati ad una rispettiva pluralità di ulteriori blocchi costituenti il macroblocco puntato  
35       dal vettore di moto trovato in precedenza, in cui detta

prima fase di identificazione viene applicata a detto macroblocco, mentre detta seconda fase viene applicata ai blocchi di detta pluralità per identificare detta pluralità di vettori.

5           18. Sistema per la stima del movimento in segnali video (10) organizzati in frame successivi suddivisi in macroblocchi tramite l'identificazione di vettori di moto, il sistema comprendendo un blocco di stima del movimento (11) configurato per attuare una prima fase  
10 di identificazione in cui, a partire da un vettore di moto corrente un predittore di vettore di moto migliore viene identificato nell'ambito di un insieme di candidati ed una seconda fase di affinamento del predittore migliore così identificato; detto insieme di  
15 candidati essendo formato a partire da vettori appartenenti a macroblocchi prossimi a detto vettore corrente nell'ambito del frame corrente e del frame precedente.

          19. Sistema secondo la rivendicazione 18,  
20 caratterizzato dal fatto che detto blocco di stima (11) è configurato per includere in detto insieme include, nel frame precedente, il vettore omologo a detto vettore di moto corrente.

          20. Sistema secondo la rivendicazione 18 o la  
25 rivendicazione 19, caratterizzato dal fatto che comprende l'operazione di sottoporre a scansione detti macroblocchi in ordine lessicografico nell'ambito di ciascun frame, per cui detto insieme include, nell'ambito del frame corrente, vettori appartenenti a  
30 macroblocchi disposti al di sopra ed a sinistra del macroblocco corrente.

          21. Sistema secondo una qualsiasi delle precedenti rivendicazioni 18 a 20, caratterizzato dal fatto che detto blocco di stima (11) include in detto insieme



vettori appartenenti ad almeno uno dei macroblocchi scelti nel gruppo costituito da:

- nel frame corrente, macroblocchi situati a sinistra ed al di sopra del macroblocco corrente, e
- 5       - nel frame precedente, il macroblocco omologo e i macroblocchi situati a sinistra del macroblocco corrente.

22. Sistema secondo una qualsiasi delle precedenti rivendicazioni 18 a 21, caratterizzato dal fatto che  
10       detto blocco di stima (11) sceglie detto predittore migliore, nell'ambito di detto insieme, quale predittore che minimizza una funzione di misura di errore residuo.

23. Sistema secondo la rivendicazione 22,  
15       caratterizzato dal fatto che detta funzione è la funzione SAD (somma delle differenze assolute).

24. Sistema secondo la rivendicazione 22, caratterizzato dal fatto che detta funzione è la funzione SSD (somma delle differenze al quadrato).

20       25. Sistema secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che detto blocco di stima (11) è configurato per attuare detta fase di affinamento formando una griglia di n punti centrati intorno alla posizione centrale puntata dal  
25       vettore di modo migliore (MV), i punti di detta griglia presentando una distanza rispetto a detto centro definita quale funzione, preferibilmente lineare, dell'errore di adattamento definito in detta prima fase di identificazione.

30       26. Sistema secondo la rivendicazione 25, caratterizzato dal fatto che detta distanza è definita come prodotto fra un coefficiente e detta funzione di errore.

27. Sistema secondo la rivendicazione 25 o la  
35       rivendicazione 26, caratterizzato dal fatto che detta

funzione di errore è la funzione SAD (somma delle differenze assolute).

28. Sistema secondo una qualsiasi delle rivendicazioni 25 a 27, caratterizzato dal fatto che  
5 detto blocco di stima (11) sceglie, come vettore di moto per il macroblocco corrente, nell'ambito di detta griglia, il vettore che minimizza un'ulteriore funzione di errore residuo.

29. Sistema secondo la rivendicazione 11,  
10 caratterizzato dal fatto che detta ulteriore funzione di errore residuo è la funzione SAD (somma delle differenze assolute).

30. Sistema secondo una qualsiasi delle rivendicazioni 18 a 29, caratterizzato dal fatto che  
15 detto blocco di stima (11) è suscettibile di amplificare selettivamente detta griglia per un fattore di amplificazione dato.

31. Sistema secondo la rivendicazione 30, caratterizzato dal fatto che detto blocco di stima (11)  
20 monitora la qualità di detto predittore in funzione di una funzione di errore data ed amplifica detta griglia quando detto predittore presenta una funzione di errore superiore a detto valore dato.

32. Sistema secondo la rivendicazione 30 o la  
25 rivendicazione 31, caratterizzato dal fatto che detto blocco di stima (11) calcola il fattore di amplificazione della griglia in funzione dell'analisi della distribuzione di detta funzione di errore per alcune sequenze campione.

33. Sistema secondo una qualsiasi delle rivendicazioni 18 a 32, configurato per operare secondo  
30 il modo di funzionamento IPB dello standard video H.236+ per il trattamento di coppie di frame PB, in cui detto blocco di stima (11) sceglie i predittori  
35 temporali del frame B di una coppia BP prima dei

predittori temporali del frame P della coppia BP stessa.

34. Sistema secondo una qualsiasi delle rivendicazioni 18 a 32, configurato per operare secondo  
5 il modo di funzionamento APM dello standard video H.263+, in cui detto blocco di stima (11) associa al frame di un macroblocco una pluralità di vettori, preferibilmente in numero di quattro, associati ad una rispettiva pluralità di ulteriori blocchi costituenti  
10 il macroblocco puntato dal vettore di moto trovato in precedenza, in cui detto blocco di stima (11) applica rispettivamente:

- detta prima fase di identificazione, a detto macroblocco, e

15 - detta seconda fase, ai blocchi di detta pluralità per identificare detta pluralità di vettori.

## RIASSUNTO

La stima del movimento in segnali video (10) organizzati in frame successivi suddivisi in macroblocchi viene svolta tramite l'identificazione di  
5 vettori di moto. In una prima fase di identificazione, a partire da un vettore di moto corrente viene identificato un predittore di vettore di moto migliore scelto nell'ambito di un insieme di candidati. Il predittore migliore così identificato viene quindi  
10 sottoposto ad una seconda fase di affinamento. Il suddetto insieme di candidati viene identificato scegliendo vettori appartenenti a macroblocchi prossimi al vettore corrente nell'ambito del frame corrente e del frame precedente. Di preferenza, la fase di  
15 affinamento comporta la definizione di una griglia di n punti centrata intorno alla posizione centrale puntata dal vettore di moto migliore ed i punti della griglia presentano una distanza rispetto a detto centro definita in funzione dell'errore di adattamento  
20 (tipicamente costituito da una funzione SAD) definita nella prima fase di identificazione. E' prevista l'applicazione ai modi di funzionamento IPB ed APM dello standard video H.263+.

(Figura 1)

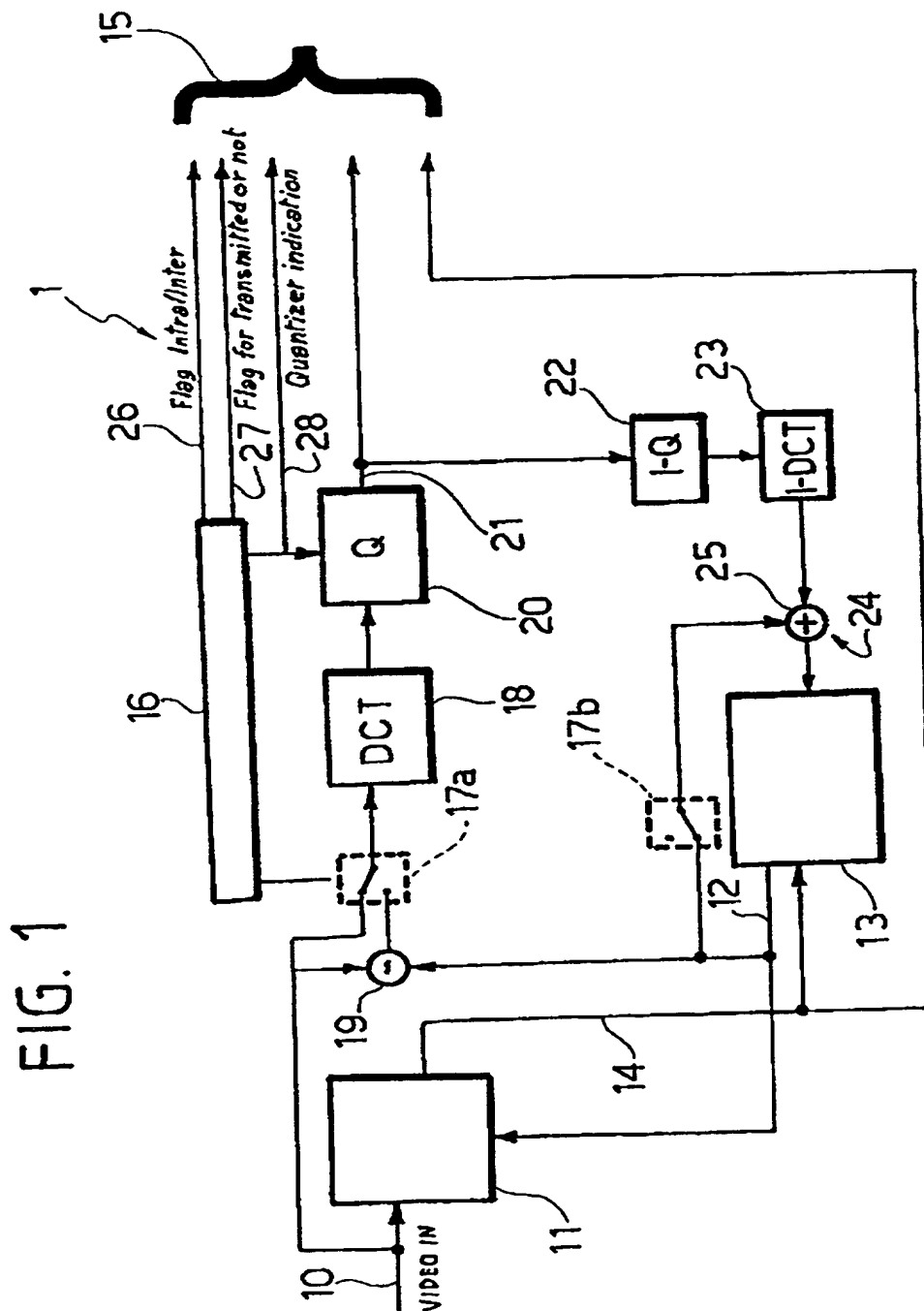


FIG. 2

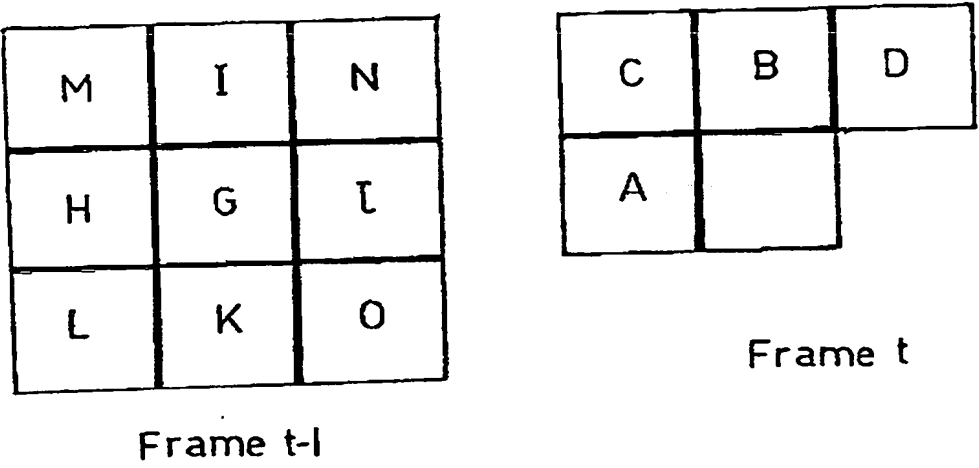


FIG. 5

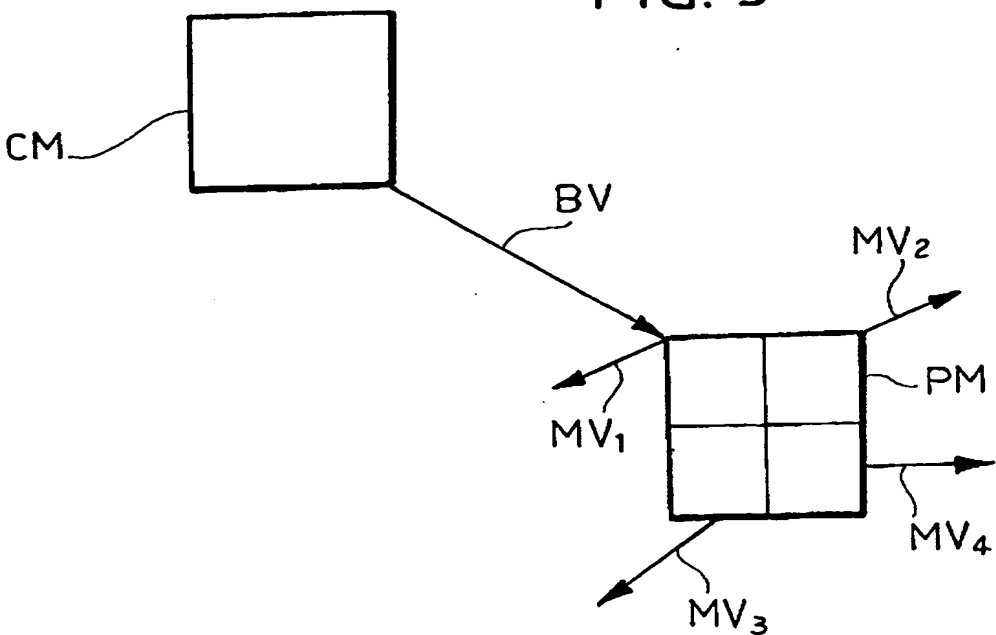


FIG. 3

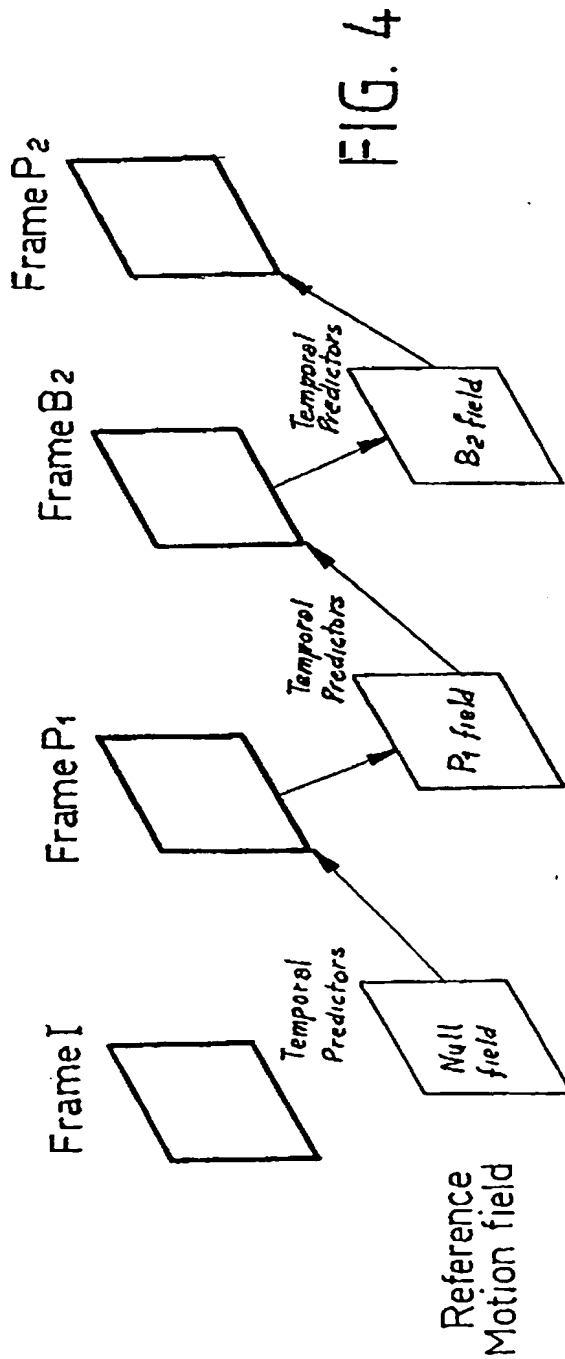
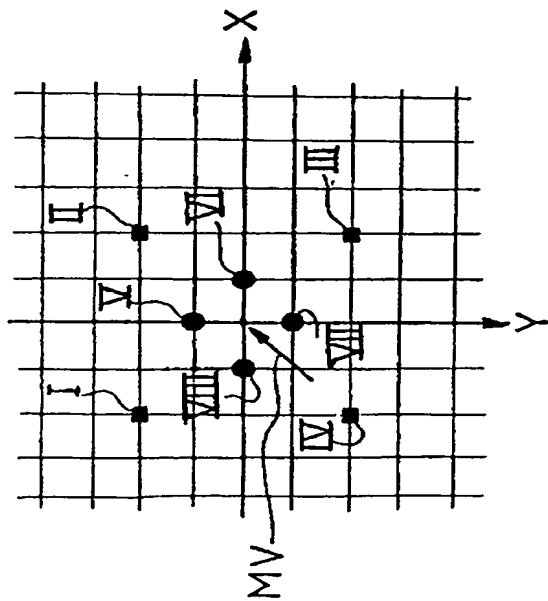


FIG. 4

**THIS PAGE BLANK (USPTO)**